

Article

Serious Games and Mathematical Fluency: A Study from the Gender Perspective in Primary Education

Fernando Fraga-Varela , Esther Vila-Couñago and Ana Rodríguez-Groba 

Department of Pedagogy and Didactics, Faculty of Educational Sciences, University of Santiago de Compostela, 15705 Santiago de Compostela, Spain; esther.vila@usc.es (E.V.-C.); ana.groba@usc.es (A.R.-G.)

* Correspondence: fernando.fraga@usc.es

Abstract: In recent years, serious games offer great opportunities for learning processes at schools. However, it is unclear whether this type of proposals can offer differentiated answers among the students according to their gender. In this context, the aim of this paper is to know the possible differences that occur in primary school classrooms according to gender, with serious games designed for the development of mathematical fluency, and to examine to what extent these games contribute to the overall school performance. We carried out a quasi-experimental study, including pretest and posttest, without control group and with several experimental groups, and the participation of 284 students from first to fourth grade. The results show that the software benefits boys and girls equally, compared to the previously followed methodology that benefited boys. A clear relation between the results achieved and the performance in the overall students' grades has also been observed. The conclusions show the potential of serious games in school settings, and the opportunity to approach performance differences based on the gender.

Keywords: gender; serious games; academic achievement; primary education; mathematics education; primary school teachers



Citation: Fraga-Varela, F.; Vila-Couñago, E.; Rodríguez-Groba, A. Serious Games and Mathematical Fluency: A Study from the Gender Perspective in Primary Education. *Sustainability* **2021**, *13*, 6586. <https://doi.org/10.3390/su13126586>

Academic Editors: Víctor Abella García, Roberto Baelo, Sheila García-Martín and Mario Grande-de-Prado

Received: 15 May 2021
Accepted: 4 June 2021
Published: 9 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The possibilities offered by serious games has been an area of increasing interest over the past years in the field of education. Unlike video games, which are created for entertainment purposes, serious games focus on the educational aspect [1] on one of the possible developments that this type of software is able to offer [2]. Serious games recover characteristic elements of the students' leisure time and take them to the classroom, generating an experience that promotes learning.

The presence of video games has progressively ousted other forms of entertainment. Data provided by the Spanish Video Game Association [3] show very high average weekly consumption figures: 6.7 h a week on average in Spain (in countries like the UK the figure almost doubles, with 11.6 h) and a male profile in 58% of the cases. This should be added to a larger presence in the lower age ranges: 75% of the boys and girls aged between 6 and 10 years fit this profile, thus reducing the percentage as we include higher ages. With these consumption levels, we find studies that point out the benefits of video games [4], also shared with serious games [5]. However, they are not exempt from possible negative assessments either. Research exists [6] showing that the use of video games during school days is related to a worse academic performance. Moreover, intensive video gamers have worse academic results in a research line that relates excess time dedicated to this type of leisure with problems in school performance [7], a situation occurring twice as often among boys than among girls [8].

However, the use of serious games offers an opportunity, because students showing difficulties working in ordinary situations at school increase their engagement in this type of proposals [9]. We do also have to be aware of possible improvements in learning processes, which are highly dependent on the contexts [10] and the design of the game experience [11].

The relation between motivation and learning benefits offers an opportunity to explore new proposals applied on teaching and learning.

The improvements produced by the use of serious games are summarized in three distinct types of engagement [12] related to behavior, affect and cognition. This explains the enormous impact that they usually have on the students' motivation [13], which results in an improvement of the participation level and high engagement levels [14]. This strong link with software may help to understand the advantages in learning processes [15,16] and school curriculum [17].

1.1. Serious Games and Gender

Research work in the field of video games from a gender perspective is often found. In this sense, several studies have been carried out, dealing with the personality of boys and girls and the relation with the type of preferred video game [18] or how gender governs the interaction between cognitive and personal factors [19]. These works help to understand the relation between gender motivations that lead boys to spend more time playing [20]. Moreover, some works have shown that in the case of women, the relation between the type of game chosen and future fields of academic interest is stronger than in the case of men [21].

As opposed to the large volume of studies on gender and video games, there are only a few works focusing on this field in serious games. The analyzed studies show differences at social interaction level [22] and investigate differences in academic performance, concluding that in the case of girls it is, unexpectedly, as good as among boys, despite completing less missions and scoring worse in the game [23]. There are even studies according to which girls exceed boys in terms of commitment to the proposal and the learning level achieved [24,25], which is evidence of the concern of women to relate the proposed experience with the academic achievement [26]. These data apparently contradict some studies in which the teachers do state gender differences in the motivational capacity perceived in favor of boys [27]. Other works [28], although they do not find differences in performance, show differentiated styles: women collaborate and are more willing to accept experiences with serious games, while men are more competitive.

1.2. Serious Games and Mathematics

In primary education schools, serious games can help to improve the learning of the subjects, and, among them, the field of mathematics is a recurrent space of interest. It is also a place of research due to the gender gaps that affect achievements, attitudes and relationship with this field of knowledge [29]. Research shows that these gaps change based on cultural variations and opportunities for girls and women in aspects such as equity in school enrolment or women's participation in research work [30]. However, at the primary education level there are no differences between boys and girls in mathematics performance, a situation that changes in favor of boys in secondary school and university [31]. What is identified in this type of content is that the gender difference in the performance in activities involving competitiveness among students is not the same as that produced in non-competitive tasks. This competitiveness, when it enters the picture, could be affecting the size of the gender gap in mathematics test scores, potentially exaggerating the mathematical advantage of men over women when learning is similar [32]. All of these elements globally influence the performance of boys and girls and explain, together with factors such as the socioeconomic level, the significant differences shown in international studies for the Spanish context compared to other countries [33].

One way to see the progress in the first levels of primary education in mathematics is through the concept of mathematical fluency [34]. The measure of this element can work as a good indicator of school improvement with the use of serious games. Recent research relates the use of serious games with the improvement of mathematical fluency [35], being also a benefit that is projected in the set of key learning in schools [36].

In order to probe into the possible improvements in mathematical fluency due to the use of serious games, we have recently carried out an investigation in this area with the ReflexMath software. It is a program frequently used in schools in the United States that focuses on the mathematical fluency of calculus (Figure 1) with a curricular structure that follows the indications of the Council of Chief State School [37]. The software generates fully individualized learning pathways from the massive collection of user data, mainly based on response times and the error rate. A virtual assistant, in the form of a child character, provides help when the calculation processes do not improve. In serious games, the activation of the assistant is based on the error, considering it an opportunity for learning and not a penalty. This character, following the indicated curricular proposal, offers specific strategies for the calculation operation detected as especially difficult for students. The work of systematizing the calculation is based on video games in which the user's interaction with the characters and the proposed situations is achieved by solving calculation operations adapted in real time to their profile. The software, in addition to the serious game, has an integrated gamification system generating rewards through exchangeable points for improvements to the game's avatar, awards in the form of diplomas, etc. similar to a microform of digital badging. Teachers have the possibility of monitoring the process through a specific dashboard that offers information through tools based on learning analytics (Figure 2).

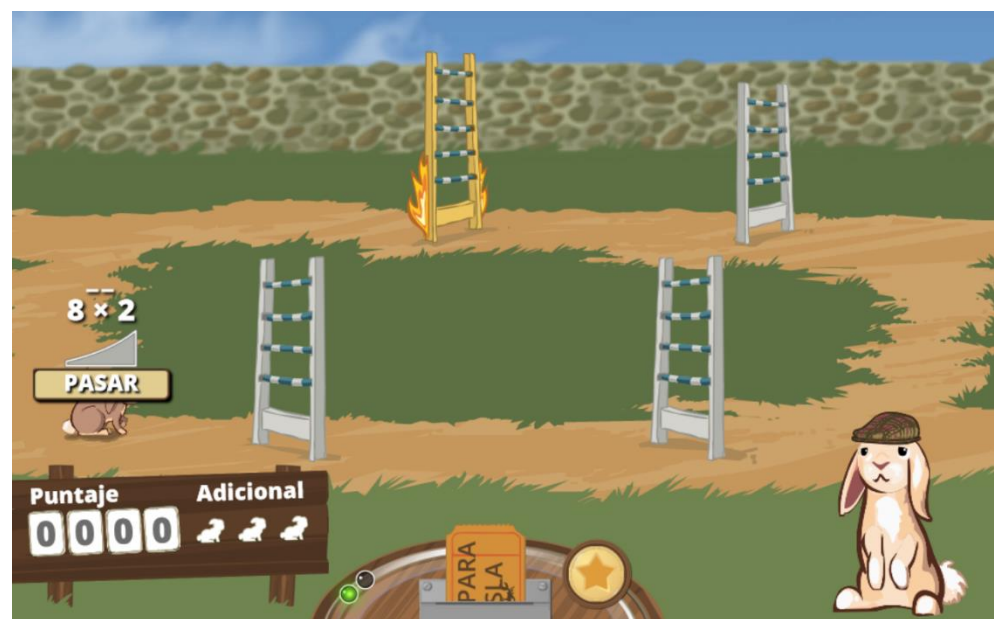


Figure 1. A game in ReflexMath software.

The data from the research carried out with this software in a primary education center show that, in a context such as Spain and with a curriculum structured differently from the American one, this software leads to great improvements in mathematical fluency in schools with ordinary schooling students in the first four levels of primary education with a statistically significant and large improvement in mathematical fluency. The data are consistent with previous studies with the same software with much smaller samples [38] and also with students with learning difficulties [39].

However, surprisingly, the differences in performance shown in the pretest segmented by gender turn out to be very favorable towards boys. The pretest data also showed that, with the methodology used in the school center, these differences did not decrease over the years. In addition, it was an aspect in which the educational center of the study showed concern because it had never been assessed and there was no awareness of this fact. It is an emerging element that requires analysis and that could potentially be worked on by means of an alternative methodology through a serious game.

For all of these reasons, the objectives of this research work are, on the one hand, to analyze from a gender perspective the performance of students in mathematical fluency before and after using a serious game, and the use that they have made of the serious game in terms of number of days they have used it and the number of activities resolved; and, on the other hand, to study from a gender perspective the role that classroom work with serious game plays in the general school performance of students.

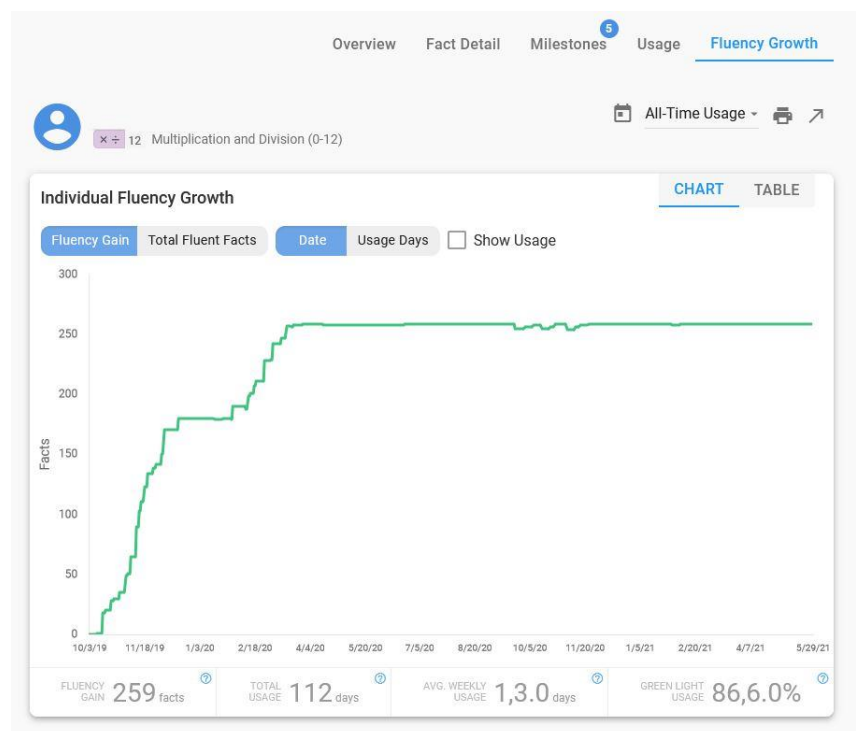


Figure 2. Teacher dashboard.

2. Materials and Methods

2.1. Design

To achieve the intended objectives, a quasi-experimental design with pretest-posttest was proposed, without a control group and with several experimental groups. The study was carried out in a standard school setting, where the classroom-groups were already established (the assignment of the subjects to the groups was not random). Each group has worked completely independently, without instructor-initiated intervention, using the ReflexMath educational software. A mathematical fluency test was applied before and after its use. The teachers of the mathematics area were in charge of putting the proposal into practice in each classroom. To do so, the same indications were given to them on the use of the possibilities offered by the system. For the performance of the study, we had the informed consent of the educational center.

2.2. Participants

In this research there were 12 primary education classrooms between the first and fourth grade levels, belonging to a private (concerted) educational center located in an urban context of the Autonomous Community of Galicia. Specifically, 284 students participated, of which 54.2% were boys and 45.8% were girls. As regards the academic year, the sample was configured as follows: 24.3% first grade; 25.4% second grade; 25.7% third grade; and 24.6% fourth grade. First, the proposal was presented to the school leadership team and the 25 teachers who teach in the primary education stage. The participation in the research work received a positive feedback, especially by the mathematics teachers.

2.3. Instrument

The assessment of calculus learning was carried out through the Basic Math Operations Task (BMOT) [40], which was extracted from a later publication [39], and which was completely translated into Spanish, both in its pretest and posttest versions. This instrument includes combined operations for calculating additions, subtractions, multiplications and divisions, which the students must answer in a maximum time of one minute. The test is entirely suited to the curricular level of third and fourth grades of primary education, while for the first and second grades it was adapted including only addition and subtraction. The test correction generates an individual performance indicator based on the total rate of operations that are answered correctly. This indicator matches the object of work of the software, justifying its relevance.

2.4. Procedure

The study was carried out between September 2019 (application of the pretest) and December 2019 (application of the posttest). The software was used during the classroom work time in the math classes, at the rate of three sessions a week, as recommended by the manufacturer. Data were collected on the gender of the students and their academic performance, specifically the grades of all the knowledge areas assessed at the end of the first quarter of the school year, coinciding with the posttest. It was also been possible to obtain information on the number of days of use of the program and the volume of activities resolved in each case, through a specific tool that allows access to the system database. Moreover, as regards the teaching staff, data was taken from the years of teaching experience.

2.5. Data Analysis

The software used for data analysis has been IBM SPSS Statistics, version 25. Regarding the first objective, measures of central tendency and dispersion were calculated and the Student's *t* test of independent samples was applied to find out if there were statistically significant differences based on gender. For the comparison between the pretest and the posttest within the same group of male or female students, the *t*-test was applied for related samples. For these statistical hypothesis tests, a significance level of 0.05 was established. In addition, to know the size of the effect, coefficients related to Cohen's *d* family were used, specifically the formulas corresponding to d_p , in the case of the independent groups (boys vs. girls), and to d_D , in the case of the related measures test-posttest within the same group [41]. The resulting values were interpreted according to conventional criteria [42]: 0.2 implies a small effect size; 0.5, medium; and 0.8, large. In relation to the second objective, simple regression analysis was used, given the nature of the criterion variable (general academic grade) and attending to a single explanatory variable (the posttest score, the number of days of use and the number of activities performed with ReflexMath were taken separately).

3. Results

First, the results in mathematical fluency and use of serious play are presented according to gender. Next, we go deeper into the different educational levels and descend to certain classrooms to analyze in detail the differences between male and female students. Finally, we investigate to what extent serious play contributes to the general school performance of students.

3.1. Mathematical Fluency in Primary Education: Differences Based on Gender

The results obtained by primary school students, segmented according to gender, show an unbalanced situation (Table 1) with a significant difference of one point and a half in favor of boys in mathematical fluency in the pretest ($t(282) = 2.153$, $p = 0.032$, $d = 0.26$). In the posttest, after the use of serious game, the differences remain ($t(282) = 2.496$, $p = 0.013$, $d = 0.30$). This situation offers revealing elements: the methodology used in the school

center favors boys in general, but serious play produces an improvement for boys and girls alike, since gender differences do not change after the intervention. However, there are also apparently contradictory data: girls develop a similar improvement to boys doing fewer activities. This difference is significant, but of low magnitude ($t(282) = 2.291$, $p = 0.023$, $d = 0.27$) and offers an advantage to girls in performance per activity performed compared to boys who do more to obtain the same level.

Table 1. Basic descriptive statistics according to the gender of the students.

	Boys ($n = 154$)			Girls ($n = 130$)		
	Mean	Median	SD	Mean	Median	SD
Pretest Score	9.68	9	6.002	8.18	8	5.709
Posttest Score	18.80	18.50	7.710	16.58	16	7.127
Days of use	27.1	26	7.4	27.3	26.5	7.1
Activities resolved	6155	5222	3555	5265.5	4519.5	2871.1

If we analyze how much improvement is achieved with the use of serious games in each group by gender, at the intragroup level, we see that both boys ($t(153) = -19.340$, $p = 0.000$, $d = 1.56$) and girls ($t(129) = -17.650$, $p = 0.000$, $d = 1.55$) make great progress thanks to the serious game between pretest and posttest, but without actually differentiating between them ($t(282) = 1.052$, $p = 0.294$, $d = 0.13$). Figure 3 reflects the distributions of the scores before and after the use of the serious game, showing the important progress, as well as a greater dispersion in the posttest.

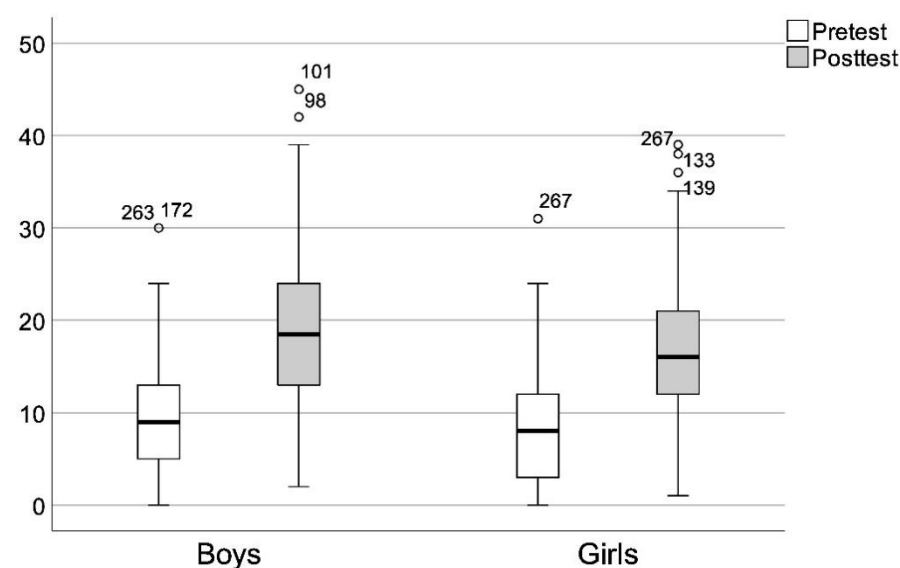


Figure 3. Distribution of pretest and posttest scores according to gender.

3.2. Mathematical Fluency, Gender and Educational Level

If the data is broken down by educational levels and by gender (Table 2), distributing the participating students with similar curricular requirements, we observe that the data by grades favor in the posttest those who had a better performance in the pretest. Between boys and girls the differences become significant in the second grade both in the pretest ($t(70) = 3.345$, $p = 0.001$, $d = 0.80$) and in the posttest ($t(70) = 2.939$, $p = 0.004$, $d = 0.71$).

Table 2. Basic descriptive statistics according to the grade and gender of the students.

1st Grade	Boys (<i>n</i> = 35)			Girls (<i>n</i> = 34)		
	Mean	Median	SD	Mean	Median	SD
Pretest Score	3.66	2	3.933	2.26	2	1.896
Posttest Score	13.37	13	6.083	12.21	13	4.670
Days of use	28.3	27	5.1	28.3	28	4.7
Activities resolved	6119.3	5139	3267.3	5192.1	4744.5	2247.3
2nd Grade	Boys (<i>n</i> = 43)			Girls (<i>n</i> = 29)		
	Mean	Median	SD	Mean	Median	SD
Pretest Score	9.74	10	3.519	7	7	3.251
Posttest Score	23.19	22	7.716	17.72	16	7.759
Days of use	30.3	30	6.7	28.7	28	5.05
Activities resolved	7472.6	6110	4282.4	5741.8	5240	3094.3
3rd Grade	Boys (<i>n</i> = 36)			Girls (<i>n</i> = 37)		
	Mean	Median	SD	Mean	Median	SD
Pretest Score	11.25	10	4.982	10.05	10	4.301
Posttest Score	15.81	15.50	5.408	15.11	15	6.050
Days of use	28.8	29	8.3	30	30	9.2
Activities resolved	7018.1	6769.5	3169.4	6233	5213	3450.3
4th Grade	Boys (<i>n</i> = 40)			Girls (<i>n</i> = 30)		
	Mean	Median	SD	Mean	Median	SD
Pretest Score	13.48	12	6.528	13.70	13	5.415
Posttest Score	21.53	21	6.857	22.27	21	6.130
Days of use	21.2	19.5	5.6	21.4	22	4.7
Activities resolved	3993.1	3411.5	2028.2	3695.1	3445	1708.8

What is evident at all levels is that girls always do fewer activities than boys. Again we find a grade, second, where the differences become statistically significant at a level of $p < 0.1$ and with an effect size very close to a mean value ($t(70) = 1.870$, $p = 0.066$, $d = 0.45$), boys developing, on average, 1731 more activities than girls.

As in the results obtained in general, between the pretest scores and those achieved in the posttest within the group of male students and the group of female students, statistically significant and highly important differences are shown in all the analyzed grades (Table 3). These data once again value the proposal with serious games compared to the work that is usually done in the school for this type of content. All students progress without gender differences, so that, fundamentally, the improvement experienced is similar for male and female students, except again for the situation in second grade at a level of $p < 0.1$ and with an effect size very close to a mean value (first grade: $t(67) = -0.216$, $p = 0.830$, $d = 0.05$; second grade: $t(70) = 1.883$, $p = 0.064$, $d = 0.45$; third grade: $t(71) = -0.441$, $p = 0.660$, $d = 0.10$; fourth grade: $t(68) = -0.485$, $p = 0.629$, $d = 0.12$).

Table 3. Statistic data from the t test for related samples (test-posttest) and size of the effect.

	Boys				Girls			
	t	df	p	d	t	df	p	d
1st grade	−12.036	34	0.000	2.03	−14.894	33	0.000	2.56
2nd grade	−15.050	42	0.000	2.30	−9.283	28	0.000	1.72
3rd grade	−6.215	35	0.000	1.04	−5.899	36	0.000	0.97
4th grade	−11.753	39	0.000	1.86	−10.391	29	0.000	1.90

3.3. Gender Differences in the Classroom: The Role of Teachers

Making a more detailed analysis of the situation in second grade, the results of the pretest-posttest show a classroom where the performance of the girls is clearly different to the rest of the groups, specifically classroom 6, as illustrated in Figure 4.

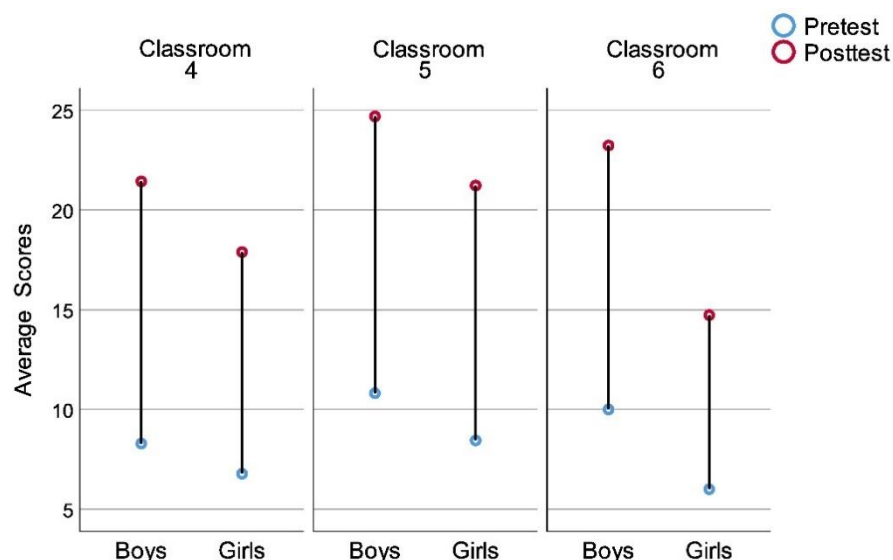


Figure 4. Progress made in mathematical fluency by the boys and girls in the three second grade classrooms.

It is a classroom where the situation of the female students in comparison with their peers (Table 4) shows statistically significant differences and of great magnitude due to gender, both in relation to the pretest ($t(22) = 3.131$, $p = 0.005$, $d = 1.28$) and the posttest ($t(22) = 3.853$, $p = 0.001$, $d = 1.58$). Questions should be raised about whether there is a different experience with serious games, but there is no data to indicate this. If we look at the time of use, both boys and girls used the software for a similar number of days and there are no statistically significant differences between them ($t(22) = 0.831$, $p = 0.415$, $d = 0.34$).

Table 4. Descriptive statistics according to the students' gender in second grade classrooms.

Classroom 6	Boys ($n = 13$)			Girls ($n = 11$)		
	Mean	Median	SD	Mean	Median	SD
Pretest Score	10	10	2.887	6	5	3.376
Posttest Score	23.23	24	5.674	14.73	14	5.022
Days of use	28.5	31	6.1	26.7	28	4.3
Activities resolved	6581.8	6110	2413.2	4715.9	4656	1434.5
Classrooms 4 and 5	Boys ($n = 30$)			Girls ($n = 18$)		
	Mean	Median	SD	Mean	Median	SD
Pretest Score	9.63	9.5	3.801	7.61	7.5	3.109
Posttest Score	23.17	21	8.538	19.56	19	8.658
Days of use	31	29.5	6.9	29.9	28.5	5.2
Activities resolved	7858.6	6078	4862.1	6368.7	5634.5	3669.1

However, in this classroom the gender gap is even accentuated, as statistically significant differences and of great magnitude are obtained in the number of solved activities ($t(22) = 2.246$, $p = 0.035$, $d = 0.92$). This situation becomes even more evident ($d = 1.05$) in the progress of mathematical fluency ($t(22) = 2.557$, $p = 0.018$). Boys improve by an average of 4.5 points more than girls.

The analysis of other parameters creates the need to reconsider the situation and opens the way to future research work. The teacher of classroom 6 is a novice teacher with an experience of less than five years and with a recent university degree. If the results obtained by the female students of this novice teacher (classroom 6) are analyzed in comparison with those of the more senior teachers (more than 30 years of experience) responsible for classrooms 4 and 5, it is observed that, despite not obtaining statistically significant differences, they are relatively important differences (mean value effect size) in relation to the pretest ($t(27) = -1.311, p = 0.201, d = 0.50$), the posttest ($t(27) = -1.678, p = 0.105, d = 0.64$), the days of use ($t(27) = -1.722, p = 0.096, d = 0.66$) and the number of activities resolved ($t(27) = -1.421, p = 0.167, d = 0.54$). By contrast, the differences are less between male students (pretest: $t(41) = 0.310, p = 0.758, d = 0.10$; posttest: $t(41) = 0.025, p = 0.980, d = 0.01$; days of use: $t(41) = -1.109, p = 0.274, d = 0.37$; activities performed: $t(41) = -0.896, p = 0.376, d = 0.30$). It would then make sense to investigate what elements could be influencing the process and how the role of the teacher could be key in the differences with other classrooms.

3.4. Contribution of Serious Games to Academic Performance

Considering the sample as a whole, the simple regression model developed for the global academic grade with the posttest score as an independent variable (Figure 5), has explained 14% of the variance (adjusted $R^2 = 0.14, F = 48.421, p = 0.000$). It is also observed that the days of use of the software (adjusted $R^2 = 0.07, F = 21.321, p = 0.000, y = 7.02 + 0.03 * x$) and the number of resolved activities (adjusted $R^2 = 0.08, F = 25.977, p = 0.000, y = 7.41 + 7.156 * 10^{-5} * x$) contribute significantly to the general academic average of the students. These data highlight the importance of these types of improvements in mathematical fluency and the potential that the use of serious games has for the development of general academic performance.

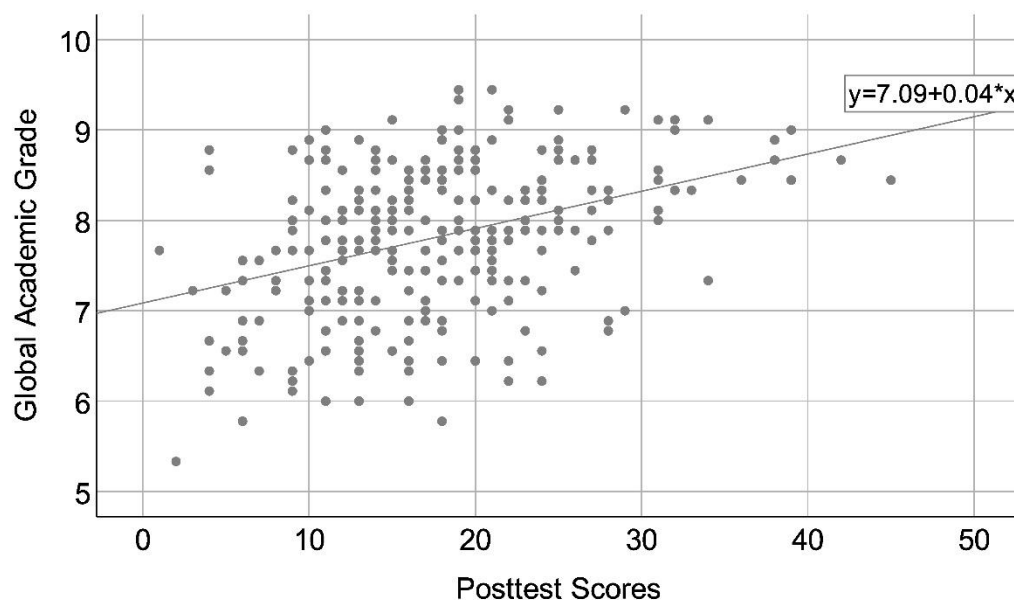


Figure 5. Scatter diagram of the posttest scores by the overall academic grade.

When calculating regression models by gender, it is noted that the posttest score explains, in the case of boys, 22% of the variance of the overall academic grade (adjusted $R^2 = 0.22, F = 43.75, p = 0.000, y = 6.75 + 0.05 * x$), while for girls this variance is reduced to 11% (adjusted $R^2 = 0.11, F = 16.466, p = 0.000, y = 7.36 + 0.04 * x$). Therefore, it is observed that mathematical fluency is, in general, a better predictor of academic performance for boys than for girls.

However, the days of use of the software and the number of activities resolved have greater explanatory power in the school performance of girls than in that of boys. The days of use manage to explain 14% of the variance of the general academic mean of the female students (adjusted $R^2 = 0.14$, $F = 22.000$, $p = 0.000$, $y = 7.36 + 0.04 * x$) and 3% in the case of male students (adjusted $R^2 = 0.03$, $F = 5.098$, $p = 0.025$, $y = 6.83 + 0.02 * x$). Likewise, the activities carried out with the software also explain 19% of the girls' performance (adjusted $R^2 = 0.19$, $F = 31.055$, $p = 0.000$, $y = 7.34 + 1.18 * 10^{-4} * x$) compared to 5% in the case of boys (adjusted $R^2 = 0.05$, $F = 9.506$, $p = 0.002$, $y = 7.35 + 5.645 * 10^{-5} * x$).

In order to better understand the role of mathematical fluency in the school performance of boys and girls, the results were broken down by academic levels (Table 5). The regression models generated were significant at a level of 0.05 both for the case of male students (first grade: $F = 22.564$, $p = 0.000$; second grade: $F = 24.272$, $p = 0.000$; third grade: $F = 4.824$, $p = 0.035$; fourth grade: $F = 18.347$, $p = 0.000$) and the female students (first grade: $F = 30.728$, $p = 0.000$; second grade: $F = 23.797$, $p = 0.000$; third grade: $F = 7.759$, $p = 0.009$; fourth grade: $F = 5.986$, $p = 0.021$). Contrary to what is seen in general, the resulting determination coefficients indicate that mathematical fluency explains the overall academic grade of girls to a greater extent than that of boys; except in fourth grade where, for male students, the proportion of the variance, explained in contrast to the lowest value of the female students, is doubled. That is, we can observe that the explanatory capacity of mathematical fluency on the school performance of boys and girls shows variations according to the analyzed grade. In any case, the estimated models have greater explanatory force in first and second than in third and fourth grade, in line with the greater curricular load of this type of content in the first two levels of primary school.

Table 5. Simple linear regression models differentiated by gender and course for the global academic grade (independent variable: posttest score).

Boys		B	SE	β	t	p	R^2	Adj. R^2
1st grade	(Constant)	6.252	0.276		22.661	0.000	0.41	0.39
	Posttest score	0.089	0.019	0.637	4.750	0.000		
2nd grade	(Constant)	6.676	0.260		25.704	0.000	0.37	0.36
	Posttest score	0.052	0.011	0.610	4.927	0.000		
3rd grade	(Constant)	7.263	0.348		20.868	0.000	0.12	0.10
	Posttest score	0.046	0.021	0.353	2.196	0.035		
4th grade	(Constant)	5.760	0.415		13.885	0.000	0.33	0.31
	Posttest score	0.079	0.018	0.571	4.283	0.000		
Girls		B	SE	β	t	p	R^2	Adj. R^2
1st grade	(Constant)	6.738	0.240		28.128	0.000	0.49	0.47
	Posttest score	0.102	0.018	0.700	5.543	0.000		
2nd grade	(Constant)	6.684	0.242		27.587	0.000	0.47	0.45
	Posttest score	0.061	0.013	0.684	4.878	0.000		
3rd grade	(Constant)	7.521	0.296		25.377	0.000	0.18	0.16
	Posttest score	0.051	0.018	0.426	2.786	0.009		
4th grade	(Constant)	6.397	0.556		11.498	0.000	0.18	0.15
	Posttest score	0.059	0.024	0.420	2.447	0.021		

4. Discussion and Conclusions

The objective of this research has focused on the analysis from a gender perspective in the improvement of mathematical fluency with serious games and their relationship with school performance. Empirical evidence on the improvement that occurs in both boys and girls is provided, in line with previous research with similar software proposals [15–17].

The analysis of the data disaggregated by gender indicates a starting situation, with the methodology used in the school center, that clearly benefits boys. However, after

working with the serious game, the data show that the improvement on the part of girls is similar to that of boys. Both improve in an equivalent way, and this is in itself a fact to be highlighted. It should also be taken into account that this process is not enough to reduce the disadvantage of the girls evidenced in their mathematical fluency in the pretest. This data is consistent with previous research in which this type of proposal benefits girls equally [24,25]. In any case, we should bear in mind the fact that the work styles are different, since in the case of girls their answer is based on their concern for academic achievement [26]. This trend, as reflected in other studies [28], implies a greater acceptance of the proposal by girls, compared to a more competitive position in the case of boys. This helps to understand the greater number of activities carried out by boys, which shows findings from previous works where gender motivation by cognitive and personal factors [19] influences their engagement in the proposal [20].

The data offer a use of the BMOT instrument [40] for a purpose not contemplated in its design: rapid identification of possible gender biases in mathematics. In response to the difference in performance observed in the pretest-posttest contrast with the use of the serious game, an opportunity is provided to normalize performance by gender, with a view to achieving results closer to those obtained in research on this type of content, since at these educational levels no gender differences should be found in their learning [31]. The data confirm the power that contexts have in the use of serious games [10], in addition to their design as a learning experience [11] in order to offer the same opportunities to both genders. In this sense, the identification of a classroom that showed a clearly inferior performance of the girls, unlike the rest of the center, brings up the importance of not losing sight of the role of the teachers and the arrangement of the elements to favor the same learning opportunities for boys and girls regardless of the software and the proposal. It is essential to train and prepare teachers so that their actions can avoid an unbalanced situation from a gender perspective. It is also necessary to offer tools for their diagnosis, since in many cases their appraisals may differ from the reality that exists in their classrooms [27]. In any case, it is convenient to watch out for the experience that the serious game offers, since the greater competitiveness of boys can offer a false appearance of learning [32] made visible here in a greater number of tasks without therefore having a better performance in his mathematical fluency.

Conversely, the regression analysis shows the importance of the proposal for the set of learning, highlighting the relevance of this type of skills to understand the general school success of students [36] at an age where the curricular weight of these contents it is much higher than in higher grades. The data here do show differences based on the indicators that are taken into consideration, apparently influenced by cognitive styles [28]. We should recall that, in both cases, both boys and girls show similar performance levels in mathematical fluency. However, in the case of girls, their general academic level is more related to the time of use of the proposal and the number of activities carried out. It appears that here too, differentiating characteristics of one and the other are projected in the styles with which they participate in the serious game [19].

Finally, it is worth considering some limitations of this research. Above all, it is noteworthy that, although there is a sufficient number of subjects, data are only available from one educational center, so the results should be used with certain caution to prevent potential generalizations. The factors that have influenced the group that has shown gender differences that even increase with the serious game are also unknown. All this offers clear lines of future research, especially in the role of teachers, not contemplated in this research, which would help to identify elements to take into account in the initial and ongoing training processes.

Author Contributions: The three authors contributed equally to all sections of this paper. All authors contributed to the research design, prepared the first draft, revised and approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study is funded by Proyecto de I + D + i Entornos digitales e identidades de género en la adolescencia (EDIGA), with reference PID2019-108221RB-I00, within the framework of the State Program for the Generation of Knowledge and Scientific and Technological Strengthening of the System of R + D + i and of the State Program of R + D + i Oriented to the Challenges of Society, financed by the Ministry of Science, Innovation and Universities of the Government of Spain.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Patient consent was waived due to the fact that the research did not involve any personal data. For the performance of this study it is enough to have informed consent of the educational center.

Data Availability Statement: All data are available on the request to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Michael, D.R.; Chen, S. *Serious Games: Games That Educate, Train and Inform*; Thomson Course Technology: Boston, MA, USA, 2006; ISBN 978-1-59200-622-9.
2. Alvarez, J.; Djaouti, D. *Serious Games: An Introduction*; Éditions Questions Théoriques: Valenciennes, France, 2012; ISBN 978-2-917131-22-0.
3. AEVI La Industria Del Videojuego En España. Anuario 2019. Available online: <http://www.aevi.org.es/web/wp-content/uploads/2020/04/AEVI-ANUARIO-2019.pdf> (accessed on 10 October 2020).
4. Granic, I.; Lobel, A.; Engels, R.C.M.E. The Benefits of Playing Video Games. *Am. Psychol.* **2014**, *69*, 66–78. [CrossRef] [PubMed]
5. Connolly, T.M.; Boyle, E.A.; MacArthur, E.; Hainey, T.; Boyle, J.M. A Systematic Literature Review of Empirical Evidence on Computer Games and Serious Games. *Comput. Educ.* **2012**, *59*, 661–686. [CrossRef]
6. Gómez-Gonzalvo, F.; Devís-Devís, J.; Molina-Alventosa, P. El tiempo de uso de los videojuegos en el rendimiento académico de los adolescentes. *Comun. Rev. Científica Comun. Educ.* **2020**, *28*, 89–99. [CrossRef]
7. Lemmens, J.S.; Hendriks, S.J.F. Addictive Online Games: Examining the Relationship Between Game Genres and Internet Gaming Disorder. *Cyberpsychology Behav. Soc. Netw.* **2016**, *19*, 270–276. [CrossRef] [PubMed]
8. Greenberg, B.S.; Sherry, J.; Lachlan, K.; Lucas, K.; Holmstrom, A. Orientations to Video Games Among Gender and Age Groups. *Simul. Gaming* **2010**, *41*, 238–259. [CrossRef]
9. Mahmoudi, H.; Koushafar, M.; Saribagloo, J.; Pashavi, G. The Effect of Computer Games on Speed, Attention and Consistency of Learning Mathematics among Students. *Procedia Soc. Behav. Sci.* **2015**, *176*, 419–424. [CrossRef]
10. Hamari, J.; Koivisto, J.; Sarsa, H. Does Gamification Work?—A Literature Review of Empirical Studies on Gamification. In Proceedings of the 2014 47th Hawaii International Conference on System Sciences, Waikoloa, HI, USA, 6–9 January 2014; pp. 3025–3034.
11. Pourabdollahian, B.; Taisch, M.; Kerga, E. Serious Games in Manufacturing Education: Evaluation of Learners' Engagement. *Procedia Comput. Sci.* **2012**, *15*, 256–265. [CrossRef]
12. Hookham, G.; Nesbitt, K. A Systematic Review of the Definition and Measurement of Engagement in Serious Games. In Proceedings of the Australasian Computer Science Week Multiconference, Association for Computing Machinery, New York, NY, USA, 29 January 2019; pp. 1–10.
13. Filella, G.; Pérez-Escoda, N.; Ros-Morente, A. Evaluación Del Programa de Educación Emocional “Happy 8-12” Para La Resolución Asertiva de Los Conflictos Entre Iguales. *Electron. J. Res. Educ. Psychol.* **2017**, *14*, 582–601. [CrossRef]
14. Wrzesien, M.; Alcañiz Raya, M. Learning in Serious Virtual Worlds: Evaluation of Learning Effectiveness and Appeal to Students in the E-Junior Project. *Comput. Educ.* **2010**, *55*, 178–187. [CrossRef]
15. Clark, D.B.; Tanner-Smith, E.E.; Killingsworth, S.S. Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis. *Rev. Educ. Res.* **2016**. [CrossRef] [PubMed]
16. Wouters, P.; van Nimwegen, C.; van Oostendorp, H.; van der Spek, E.D. A Meta-Analysis of the Cognitive and Motivational Effects of Serious Games. *J. Educ. Psychol.* **2013**, *105*, 249–265. [CrossRef]
17. de Carvalho, C.V.; Rodríguez, M.C.; Nistal, M.L.; Hromin, M.; Bianchi, A.; Heidmann, O.; Tsalapatas, H.; Metin, A. Using Video Games to Promote Engineering Careers. *Int. J. Eng. Educ.* **2018**, *34*, 388–399.
18. Homer, B.D.; Hayward, E.O.; Frye, J.; Plass, J.L. Gender and Player Characteristics in Video Game Play of Preadolescents. *Comput. Hum. Behav.* **2012**, *28*, 1782–1789. [CrossRef]
19. Liu, C.-C. Understanding Player Behavior in Online Games: The Role of Gender. *Technol. Forecast. Soc. Change* **2016**, *111*, 265–274. [CrossRef]
20. Brooks, F.M.; Chester, K.L.; Smeeton, N.C.; Spencer, N.H. Video Gaming in Adolescence: Factors Associated with Leisure Time Use. *J. Youth Stud.* **2016**, *19*, 36–54. [CrossRef]

21. Giammarco, E.A.; Schneider, T.J.; Carswell, J.J.; Knipe, W.S. Video Game Preferences and Their Relation to Career Interests. *Personal. Individ. Differ.* **2015**, *73*, 98–104. [\[CrossRef\]](#)
22. Mansour, S.; El-Said, M. The Relationship between Educational Serious Games, Gender, and Students' Social Interaction. *WSEAS Trans. Comput.* **2008**, *7*, 640–649.
23. Nietfeld, J.L. Predicting Transfer from a Game-Based Learning Environment. *Comput. Educ.* **2020**, *146*, 103780. [\[CrossRef\]](#)
24. Khan, A.; Ahmad, F.H.; Malik, M.M. Use of Digital Game Based Learning and Gamification in Secondary School Science: The Effect on Student Engagement, Learning and Gender Difference. *Educ. Inf. Technol.* **2017**, *22*, 2767–2804. [\[CrossRef\]](#)
25. Zhao, D.; Muntean, C.H.; Chis, A.E.; Muntean, G.-M. Learner Attitude, Educational Background, and Gender Influence on Knowledge Gain in a Serious Games-Enhanced Programming Course. *IEEE Trans. Educ.* **2021**, 1–9. [\[CrossRef\]](#)
26. Vate-U-Lan, P. *Oxymoron of Serious Games in ELearning: Gender Differences from an Internet-Based Survey in Thailand*; Czech Technical University in Prague: Budapest, Hungary, 2017; pp. 6–17.
27. Passarelli, M.; Dagnino, F.; Earp, J.; Manganello, F.; Persico, D.; Pozzi, F.; Bailey, C.; Perrotta, C.; Buijtenweg, T.; Haggis, M. *Educational Games as a Motivational Tool: Considerations on Their Potential and Limitations*; SciTePress: Heraklion, Greece, 2021; pp. 330–337.
28. Garber, L.L.; Hyatt, E.M.; Boya, Ü.Ö. Gender Differences in Learning Preferences among Participants of Serious Business Games. *Int. J. Manag. Educ.* **2017**, *15*, 11–29. [\[CrossRef\]](#)
29. Alordiah, C.O.; Akpadaka, G.; Oviogbodu, C.O. The Influence of Gender, School Location and Socio-Economic Status on Students' Academic Achievement in Mathematics. *J. Educ. Pract.* **2015**, *6*, 130–136.
30. Else-Quest, N.M.; Hyde, J.S.; Linn, M.C. Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis. *Psychol. Bull.* **2010**, *136*, 103–127. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Hyde, J.S.; Fennema, E.; Lamon, S.J. Gender Differences in Mathematics Performance: A Meta-Analysis. *Psychol. Bull.* **1990**, *107*, 139–155. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Niederle, M.; Vesterlund, L. Explaining the Gender Gap in Math Test Scores: The Role of Competition. *J. Econ. Perspect.* **2010**, *24*, 129–144. [\[CrossRef\]](#)
33. TIMMS Estudio Internacional de Tendencias En Matemáticas y Ciencias. Informe Español. Available online: https://sede.educacion.gob.es/publivena/descarga.action?f_codigo_agc=21925 (accessed on 12 February 2021).
34. Baroody, A.J.; Eiland, M.D.; Purpura, D.J.; Reid, E.E. Can Computer-Assisted Discovery Learning Foster First Graders' Fluency with the Most Basic Addition Combinations? *Am. Educ. Res. J.* **2013**, *50*, 533–573. [\[CrossRef\]](#)
35. van der Ven, F.; Segers, E.; Takashima, A.; Verhoeven, L. Effects of a Tablet Game Intervention on Simple Addition and Subtraction Fluency in First Graders. *Comput. Hum. Behav.* **2017**, *72*, 200–207. [\[CrossRef\]](#)
36. Geary, D.C. Cognitive Predictors of Achievement Growth in Mathematics: A 5-Year Longitudinal Study. *Dev. Psychol.* **2011**, *47*, 1539–1552. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Council of Chief State School Officers Common Core State Standards Initiative. Preparing America's Students for College & Career. Available online: <http://www.corestandards.org/> (accessed on 19 February 2021).
38. Cozad, L. *Effects of a Digital Mathematics Fluency Program on the Fluency and Generalization of Learners*; The Pennsylvania State University: Centre, PA, USA, 2019.
39. Sarrell, D. *The Effects of Reflex Math as a Response to Intervention Strategy to Improve Math Automaticity among Male and Female At-Risk Middle School Students*; Liberty University: Lynchburg, VA, USA, 2014.
40. Foegen, A.; Deno, S.L. Identifying Growth Indicators for Low-Achieving Students in Middle School Mathematics. *J. Spec. Educ.* **2001**, *35*, 4–16. [\[CrossRef\]](#)
41. Goulet-Pelletier, J.-C.; Cousineau, D. A Review of Effect Sizes and Their Confidence Intervals, Part I: The Cohen's d Family. *Quant. Methods Psychol.* **2018**, *14*, 242–265. [\[CrossRef\]](#)
42. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; L. Erlbaum Associates: Hillsdale, NJ, USA, 1988; ISBN 978-0-8058-0283-2.